Scand. J. Lab. Anim. Sci. No.3. 1999. Vol. 26

Environmental enrichment may alter the number of rats needed to achieve statistical significance

by *Satu Eskola*[1], *Marja Lauhikari*[1], *Hanna-Marja Voipio*[2], *Matti Laitinen*[3] *and Timo Nevalainen*[1]
[1]National Laboratory Animal Center (NLAC), [3]Department of Clinical Chemistry, University of Kuopio and
[2]Laboratory Animal Centre, University of Oulu, Finland
Correspondence: Satu Eskola, NLAC, University of Kuopio, P.O.Box 1627, 70211 Kuopio, Finland

Introduction

Variation is an enemy of scientists. Less extensive variation in results simply means results achieving more significance and vice versa. Test systems and procedures are subjected to experimental variance, which can be considered under four main categories: biological, pre-analytical, analytical and pharmacological variance *(Davies, 1998)*. In attempts to control these factors, scientists have demanded standardisation of animals, procedures and assays, which ultimately means a reduction of animals used as well.

Reduction in the number of animals used is a target of the scientific community. The concept is not new - it was introduced already over 160 years ago by Marshall Hall (*Paton, 1984*) and again in the 1950's (*Russell & Burch, 1959*). With proper design of experiments, which have a high probability of achieving correct results, the number of animals used can be reduced (*Festing, 1993*). However, if too few animals are used, the experiment may fail to detect real treatment effects (*Festing, 1993; Beynen et al., 1993*), and the results are called false negatives. Hence, the use of either too few or too many animals is ethically wrong (*Erb, 1990*). A key point in designing good experiments is to control the variability of experiment (*Festing, 1994*) - both in animals and procedures.

After years of standardisation of housing, a new dimension in animal husbandry, *i.e.* environmental enrichment, is a current topic in laboratory animal science. So far studies on environmental enrichment have centred on the detection of differences in group means attributable to enrichment. For example, cage design may affect emotionality (*Chamove, 1989*) and aggression in mice (*McGregor & Ayling, 1990*), isolation may improve sexual performance in male rats, and enriched environment may retard vaginal opening in female rats (*Swanson & van de Poll, 1983*). Furthermore, the actual amount of time rats spend awake may decrease in an enriched environment (*Batchelor, 1994*). However, most environmental enrichment programmes have been introduced in order to improve the welfare of animals and not to study the effects of enrichment. Since the enrichment set-ups should be the same for all of the groups, it would not be expected to change the group means.

In many cases, the effects of enrichment programmes or items *e.g.* on physiology and behaviour of animals are subtle and detection of differences between groups is difficult. The reason for this is not necessarily the lack of significant biological effect but rather the increased variation within subjects. Indeed, these unexpected effects on the variation and consequently on the number of animals needed should be a real concern to scientists.

The aim of this study was to evaluate whether two enrichment items, an aspen tube and block have any impact on commonly analysed clinical chemistry and physiological parameters of Wistar rats and more importantly on the extent of variation. Based on the variation detected, this study went on to calculate the smallest number of animals needed to detect an arbitrarily chosen treatment effect with and without these enrichment items.

Materials and Methods

Animals and environment

A total of 48 barrier bred, but conventionally housed outbred Wistar (WH, Hannover origin) rats (National Laboratory Animal Center, Kuopio, Finland) were used. Animals were chosen from eight litters, three females and three males from each. Animals were allocated into three groups at weaning; control, tube and block group (8 males and 8 females in each). The rats were housed in stainless steel solid bottom cages (48 x 28 x 20 cm with a wire lid) in groups of four (n=4 cages per group).

The rats were three weeks old at the beginning and eight weeks old at the end of the study. The bedding used (1.2 l per cage) was aspen chips (4HP, Tapvei Oy, Kaavi, Finland), changed twice a week on Mondays and Fridays. The animals were housed at an ambient temperature of 20+2 °C and relative humidity of 42-72 %. The light/dark cycle of the animal room was 12:12 hours with lights on at 7:00. Commercial diet (R36, Lactamin AB, Södertälje, Sweden) and tap water in polycarbonate bottles were available ad libitum.

Enrichment items

In addition to the bedding material, two different enrichment items were used - a block (6x6x6 cm with penetrating drilled holes, diameter of 1.9 cm on each side) and a rectangular tube (20x12x12 cm with 1.5 cm wall thickness) made of dried aspen board. The walls of the tube were pinned together with aspen pins in predrilled holes *i.e.* no glue was used. The size of the tube was chosen so that even large rats could enter the tube, and there would still be enough space to move. The size of the block was based on an earlier study (*Kaliste-Korhonen et al., 1995*), and chosen so that it would last for at least one week if the animals gnawed it. The shape was modified from the study by Chmiel & Noonan (1996). The enrichment items were placed into the cages on the second day of the study. The items were replaced with new ones once every week on Wednesdays.

Physiological measurements

The growth of animals was followed by weighing them at the age of three, seven and eight weeks. The total growth was calculated by deducting the body weight at the age of three weeks from the final body weight (FBW).

At the age of eight weeks, the rats were sampled on two successive days between 10:00 - 16:00 h during the first day and between 9:00 - 15:00 h during the second day in a necropsy room. They were placed into an euthanasia chamber with about 70 % C_{O2} concentration, in order to achieve a rapid loss of consciousness (*Blackshaw et al., 1988*) and then maintained under constant $_{O2}$:C_{O2}-flow (1:1) with a face mask. Blood was withdrawn by cardiac puncture into gel serum vacuum tubes. The death of the animal was ensured by cervical dislocation.

Half of the animals from each group were euthanised during the first day (six cages) and the second half during the next day. Males were euthanised before females. One cage at a time was brought into the necropsy room. Final body weights as well as the weights of adrenal glands and interscapular brown adipose tissues (BAT) were measured.

The serum was separated after centrifuging at 2000 x g for 10 minutes and the sera were stored at -20 °C. The serum levels of Pi, Ca, cholesterol, triglycerides, creatinine, total bilirubin and protein were analysed with Kone Pro analyser® (Kone Instruments Corporation, Espoo, Finland) utilising Kone reagents® (Kone Diagnostics, Espoo, Finland). AFOS (alkaline phosphatase) and LDH (lactate dehydrogenase) were assessed at 37 °C according to the recommendation of Scandinavian Society of Clinical Chemistry. ALAT (alanine aminotransferase), ASAT (aspartate aminotransferase) and GGT (gamma-glutamyltransferase) were determined also at 37 °C according to the recommendation of IFCC (International Federation of Clinical Chemistry). The serum corticosterone levels were assayed with a radioimmunoassay kit (ICN Bio-chemicals, Costa Mesa, CA, USA). Seven corticosterone

samples were discarded due to an analytical error or questionably low corticosterone level (<10 ng/ml). Two samples were discarded from the other clinical chemistry analyses due to visible haemolysis, since the presence of haemoglobin in serum can cause erroneously increased values *e.g.* for LDH, ASAT and ALAT and erroneously decreased values for bilirubin and AFOS (*Leard et al., 1990; Sonntag, 1986*).

Statistical analysis

The data were processed with SPSS for Windows statistical package (Release 6.1.4, SPSS Inc., Chicago, IL, USA). The normal distribution of the data was tested with Kolmogorov-Smirnov test. Before analysing the effects of enrichment items, the effects of gender and litter were analysed. Multivariate analysis of variance test was used for normally distributed data with homogenous variance between groups and Mann-Whitney U-test (gender difference) or Kruskal-Wallis 1-way anova test (litter difference) were used for data not normally distributed or for data with a normal distribution but heterogeneous variance between groups. The weights of adrenal glands and BAT were analysed as relative weights (organ weight / FBW). Results are expressed as means ± standard deviation (SD).

Repeated measures analysis of variance test was used to analyse the effects of enrichment group, gender and litter. Since analysis of variance with litter, enrichment group and gender as independent variables could not be done (n=1 in the cells of the combinations of these three independent variables), the assumption was that the individual values for the three enrichment groups were repeated samples from the same animal, which was actually the litter number. This analysis of variance took into account the effect of litter in the analysis of enrichment and gender effects. The differences between groups were further analysed with dependent Student's t-test and the p-values were corrected with Bonferroni's formula.

Based on the means ± SDs, SOLO Power Analysis; one-sample mean (1991) was used to estimate the number of animals needed (n) to detect an arbitrarily chosen 20 % difference in the group means of each parameter, when significance was set at p=0.05 and statistical power at 0.90. The smallest accepted value for n was two. In this analysis, the BAT and adrenal weights were absolute organ weights. Additionally, so called N-ratio ($n_{enrichment}$ / $n_{control}$) was calculated (based on SOLO results) to indicate, how many times more (or less, if N-ratio is under 1) animals were needed in the enrichment group in comparison to the control group. Since n-value is based on relation of mean and SD, n-value increases when SD becomes wider or when biological effect becomes smaller (more animals are needed to detect the smaller effect) and contrary. If the N-ratio is one, the n in the enrichment group equals the n in the control group.

Results

The use of the enrichment items has been presented earlier (*Eskola et al., 1999*). Briefly, the rats spent over 80 % of their time during the light period inside the tube and about 11 % of their time during the dark period on top of the block or in its vicinity. Rats also gnawed them and the amount gnawed was essentially the same with both items and increased slightly with time.

Effects of gender and litter

Both gender and litter had effects on the physiological and serum parameters measured in the eight weeks old rats (Table 1). Significant interaction of gender and litter in AFOS activities (p=0.009) indicates that in certain litters, females and males differed from each others. No other interactions were found.

In general, male rats had higher serum AFOS (541 ± 114 vs. 283 ± 51 U/l) and ALAT activities (58 ± 11 vs. 47 ± 8 U/l) than females. Serum creatinine (44 ± 2 vs. 47 ± 3 µmol/l), protein (61 ± 3 vs. 68 ± 3 g/l) corticosterone (279 ± 214 vs. 529 ± 436 ng/ml) and LDH levels (517 ± 207 vs. 808 ± 492 U/l) were lower in male rats than in female rats. Males also had smaller relative adrenal weights (0.3 ± 0.05 vs.0.2 ± 0.04 mg/g) and they grew more (174 ± 17 vs. 124 ± 12 g) than females.

Table 1. The effects of gender and litter on the physiological and serum parameters measured in eight weeks old Wistar rats. M-W=Mann-Whitney U-test, K-W=Kruskal-Wallis 1-way anova. The statistically significant effects are in **bold**.

VARIABLES	MEAN ± SD	GENDER x LITTER F (DF), p	GENDER F (DF), p	LITTER F (DF), p
Normally distributed data and homogenous variance between groups (Multivariate analysis of variance)				
AFOS (U/l)	411.7 ± 156.7 (n=46)	**F=3.37 (7;30) p=0.009**	**F=212.48 (1;30) p=0.000**	**F=5.94 (7;30) p=0.000**
ALAT (U/l)	52.8 ± 10.7 (n=46)	F=0.92 (7;30) p=0.505	**F= 17.27 (1;30) p=0.000**	F=1.26 (7;30) p=0.303
Pi (mmol/l)	3.2 ± 0.3 (n=46)	F=1.35 (7;30) p=0.263	F=0.17 (1;30) p=0.679	**F=5.87 (7;30) p=0.000**
Cholesterol (mmol/l)	2.5 ± 0.4 (n=46)	F=1.88 (7;30) p=0.109	F=0.73 (1;30) p=0.401	**F=7.38 (7;30) p=0.000**
Creatinine (μmol/l)	45.5 ± 2.9 (n=46)	F=0.50 (7;30) p=0.830	**F= 9.15 (1;30) p=0.005**	F=2.03 (7;30) p=0.084
Protein (g/l)	64.4 ± 4.2 (n=46)	F=1.44 (7;30) p=0.226	**F= 57.26 (1;30) p=0.000**	F=1.13 (7;30) p=0.370
Relative brown adipose tissue weight (mg/g)	1.69 ± 0.33 (n=48)	F=0.58 (7;32) p=0.771	F=2.40 (1;32) p=0.131	**F=3.76 (7;32) p=0.004**
Normally distributed data but not homogenous variance between groups				
Relative adrenal weight (mg/g)	0.275 ± 0.079 (n=48)		**U= 21.0 p=0.000 (M-W)**	X^2=7.81, p=0.348(K-W)
Growth (g)	148.8 ± 29.2 (n=48)		**U=5.0 p=0.000 (M-W)**	X^2=5.39 p=0.612 (K-W)
Not normally distributed data				
Corticosterone (ng/ml)	413.7 ± 369.3 (n=39)		**U=114.0 p=0.035 (M-W)**	**X^2=16.59 p=0.020 (K-W)**
LDH (U/l)	662.3 ± 400.8 (n=46)		**U=168.0 p=0.034 (M-W)**	X^2=7.01 p=0.428 (K-W)

Two out of eight litters had the mean serum AFOS activities around 500 U/l whereas the others mainly were about 350 U/l. These two litters were probably responsible for the statistically significant difference (p=0.000). Serum Pi and cholesterol levels were similar in both genders, but there were differences between litters (p=0.000, respectively). There were also some litter differences in corticosterone levels (p=0.020). Furthermore, three litters had mean relative BAT weights over 1.7 mg/g and one litter had about 1.2 mg/g, which probably caused the statistically significant difference (p=0.004).

Effects of enrichment items, gender and litter

The effect of enrichment items in addition to the gender and litter was taken into account by using repeated measures analysis of variance (Table 2). There were no differences between genders in different enrichment groups in any of the parameters measured (gender and enrichment interactions p>0.05, data not shown).

The presence of enrichment items had an effect only on the serum total bilirubin levels (p=0.023).

Table 2. The effects of gender and enrichment items on the physiological and serum parameters measured in eight weeks old Wistar rats. The effect of litter was taken into account by using repeated measures analysis of variance. Statistically significant effects are in **bold**.

VARIABLES	MEAN ± SD	GENDER F (DF), p	ENRICHMENT F (DF), p
Relative adrenal weight (mg/g)	0.275 ± 0.079 (n=48)	**F=148.11 (1;7) p=0.000**	F=0.47 (2;14) p=0.635
Growth (g)	148.8 ± 29.2 (n=48)	**F=208.03 (1;7) p=0.000**	F=0.01 (2;14) p=0.993
AFOS (U/l)	411.7 ± 156.7 (n=46)	**F=52.95 (1;5) p=0.001**	F=1.52 (2;10) p=0.266
ALAT (U/l)	52.8 ± 10.7 (n=46)	**F=10.15 (1;5) p=0.024**	F=0.24 (2;10) p=0.788
Pi (mmol/l)	3.2 ± 0.3 (n=46)	**F=18.85 (1;5) p=0.007**	F=1.27 (2;10) p=0.322
Creatinine (µmol/l)	45.5 ± 2.9 (n=46)	**F=9.57 (1;5) p=0.027**	F=1.04 (2;10) p=0.388
Protein (g/l)	64.4 ± 4.2 (n=46)	**F=33.08 (1;5) p=0.002**	F=0.77 (2;10) p=0.491
Tot. bilirubin (µmol/l)	1.8 ± 0.9 (n=46)	F=0.62 (1;5) p=0.465	**F=5.59 (2;10) p=0.023**
Corticosterone (ng/ml)	413.7 ± 369.3 (n=39)	**F=13.22 (1;3) p=0.036**	F=2.10 (2;6) p=0.204

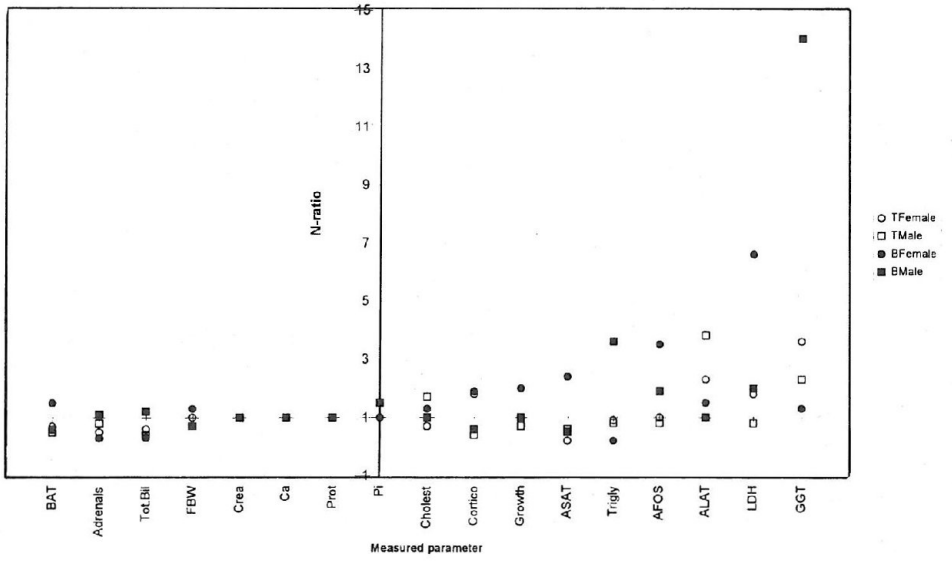
Animals with blocks (1.4 ± 0.63 µmol/l) seemed to have lower serum total bilirubin levels than animals with tubes (2.0 ± 0.76 µmol/l) or control animals (1.9 ±1.06 µmol/l). However, the post hoc-analyses did not reveal any further statistically significant differences (data not shown).

The gender differences in this statistical analysis were rather similar with the results achieved without the enrichment variable (Table 1). The differences between genders remained in the relative adrenal weights, total growth, serum AFOS, ALAT, corticosterone, creatinine and protein levels (Table 2). However, now the serum LDH levels did not reach the statistically significant difference between genders, but it was found in Pi levels (p=0.007).

Enrichment groups and N-ratios

The number of animals needed to detect 20 % difference in GGT levels was 14 times higher (N-ratio = 14) in male rats with blocks than in control males (Fig 1). The n-value of females with blocks was almost same as the n of controls (N-ratio = 1.3). The effect of tubes was somewhere in between in both genders (N-ratio in females = 3.6 and in males = 2.3). Female rats with blocks had the highest N-ratio (= 6.6) in LDH and males with tubes had the highest N-index in ALAT (= 3.8).

Fig 1. SOLO Power Analysis was used to calculate the smallest number of animals needed (n) to detect an arbitrarily chosen 20 % difference in mean of each physiological and clinical chemistry parameter, when significance was set at p=0.05 and statistical power at 0.90. Additional N-ratio of these parameters was calculated ($n_{enrichment} / n_{control}$) for female and male Wistar rats with aspen tube (T) or aspen block (B), n=8 animals per group. The parameters on the right side of Y-axis have at least one N-ratio greater than 1.5.



The presence of tubes or blocks in cages did not have any effects on the n-values of creatinine, Ca, and protein (N-ratios = 1) in either sex and minor effect on Pi in males with enrichment items (N-ratios = 1.5). Enrichment items mostly decreased the n-values in adrenal glands (N-ratios = 0.3 to 1.1) and in serum total bilirubin levels (N-ratios = 0.3 to 1.2) of both sexes. The N-ratios were also equal or under one in BAT and FBW- parameters, except for females with blocks (N-ratios = 1.5 and 1.3). The other parameters (cholesterol, corticosterone, growth, triglycerides, AFOS and ASAT) were to a greater or lesser extent affected by the presence of enrichment items (N-ratios = 0.2 to 3.

Number of animals needed by litters

The number of animals needed to detect 20 % difference in means for each litter, separately for both sexes is presented in Figure 2. The parameters with n-values smaller than six are omitted from the figure (Ca, cholesterol, creatinine, Pi, protein). In addition to former parameters, the number of animals needed in each litter was virtually the same in ALAT, AFOS, growth, FBW, BAT and adrenal parameters. The serum ASAT, GGT, triglycerides, LDH, corticosterone and total bilirubin levels were more sensitive to the effect of litter *i.e.* to genetic factors.

Each litter consisted of six animals (three females and three males), which were divided into three groups (tube, block and control), *i.e.* males (or females) in each litter were from these three different groups. Thus possible variation caused by enrichment items already affected the variation inside each litter, in addition to individual differences. Based on the above facts, it can be deduced from Figure 2 that neither litter nor enrichment items had effects on the variation of Ca, cholesterol, creatinine, Pi, protein, growth, FBW, ALAT, BAT, AFOS and adrenal weights. In other parameters (ASAT, GGT, LDH, triglycerides, corticosterone and total bilirubin), enrichment items increased within group variation, since the number of animals needed was increased. However, also litter had some effects on the n-values. If litter had influenced the number of animals needed, all the parameters would be situated at an equal level on the Y-axis. However, the n-values would still be at a higher level, since the enrichment items within litters increased the variation and hence the number of animals needed.

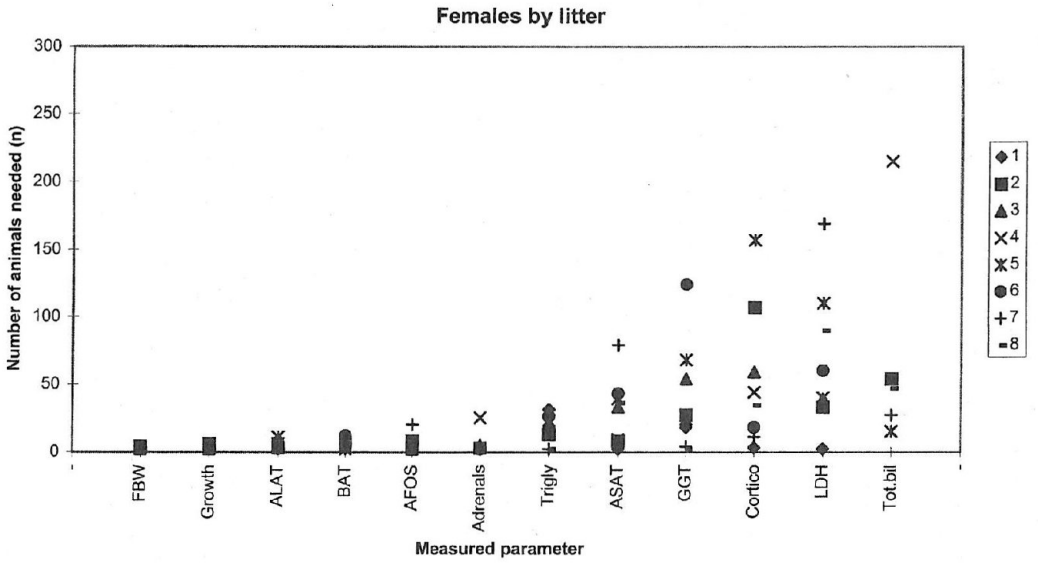
Discussion

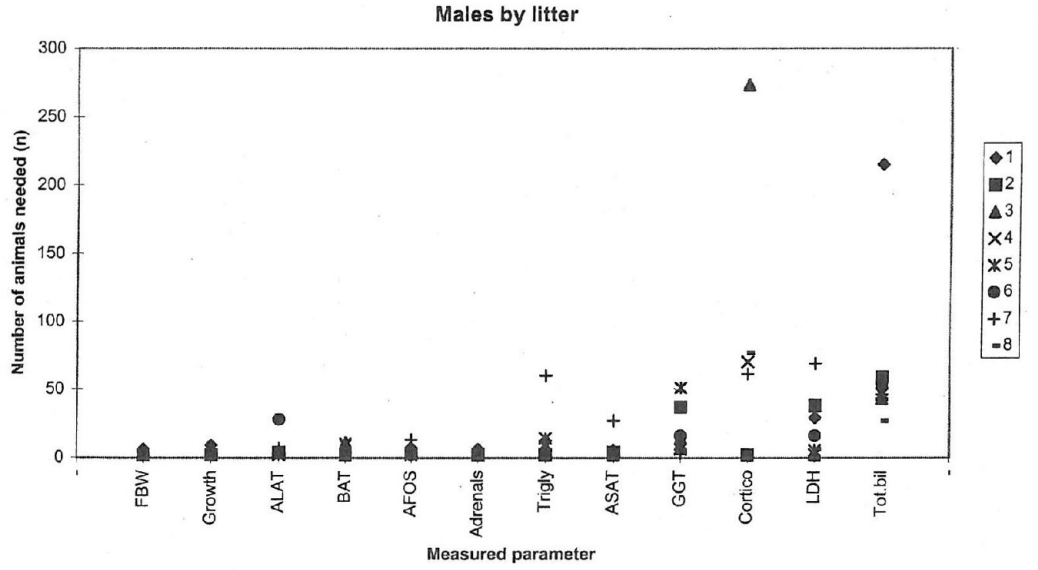
The size of an experiment can be reduced without loss of information if the variability present in material can be reduced (*Festing, 1993*). Factors causing variation in study results are either internal (*e.g.* genotype and microbiological status of animals) or external (treatment and analytical procedures). The within group and between group variation can be reduced *e.g.* by choosing more uniform animals and by standardising housing environments and treatment procedures *(Festing, 1993*). It has been suggested that enrichment can lead to a lack of standardisation, while other studies claim that it can lead to decreased variability and improved data due to reduced variation in the degree of stress (*Hubrecht, 1997*).

In the present study, we were interested in identifying which factors contributed to the variation. Those factors directly related to the animals (*e.g.* gender, age, weight, stock) and to housing environment (nutrition, light cycle, humidity, caging *etc.*) were standardised. The pre-analytical variation of physiological and serum samples was controlled as far as possible through uniform handling of animals, sample collection, separation and storage. Despite these precautions, some fluctuation was present in sample collection times; the time schedule of first day for males was 10:00 - 13:00 h and for females 14:30 - 16:00 h and that of second day for males 9:00 - 11:00 h and for females 13:00 - 15:00 h. The effects of circadian rhythm on hormones and other bodily functions are well known (*Davies, 1998*), but it is unlikely that they would have any confounding impact on these results, since the sampling times for both sexes were rather similar.

SOLO Power Analysis seems quite a useful tool for evaluating the effects of different procedures on variation. However, the calculations are based on means and standard deviations and if the data is not normally distributed these statistics are not the best describers. Furthermore, if the enrichment

Fig 2. The number of animals needed (n) to detect an arbitrarily chosen 20 % difference in mean of each physiological and clinical chemistry parameter (when significance was set at p=0.05 and statistical power at 0.90) by litter, separately for female and male rats. Each litter consists of three females and three males with either aspen tube, aspen block or without enrichment item (controls). The parameters with smaller n-value than six are omitted (Ca, cholesterol, creatinine, Pi and protein).





items would have had effects on the group means, the relation of mean and SD would have been different in each group. Therefore the n-value would have already been affected by enrichment even though the absolute SDs remained similar. In the present study most of the parameters were normally distributed within groups and the enrichment items did not really have effects (except on serum total bilirubin) on group means, thus SOLO Power Analysis should give reliable results.

The results of the present study indicate that some of the measured parameters are sensitive to the effects of enrichment. Enrichment items seemed to increase the within group variation mainly in enzyme parameters. Some of the serum parameters, *e.g.* creatinine, Ca and protein were insensitive to enrichment and consequently the number of animals needed was the same as in the control group. Furthermore, the variation within physiological parameters (BAT, adrenals and growth) was increased, decreased or unchanged depending on the enrichment item and gender. There is no straightforward explanation for this.

However, the presence of haemoglobin in serum can cause erroneously increased values *e.g.* for LDH, ASAT and ALAT and erroneously decreased values for bilirubin, AFOS and GGT (*Leard et al., 1990; Sonntag, 1986*). In the present study, visible haemolysis was detected in two samples. When these two samples were discarded from statistical analyses, the SDs for AFOS, ASAT and LDH levels decreased between 28-92 % (mean 54 %). Confusingly, the two samples with haemolysis had the highest values in AFOS (*Sonntag, 1986*). Enzymatic parameters may be more sensitive to pre-analytical and analytical variation than simple organ weighing and thus the high N-ratio values for these measures are not probably solely due to the presence of the enrichment items.

The within group variation also differed between litters, which had effects on the estimated number of animals needed. Since the number of females or males in each litter was only three, no far-reaching conclusions can be drawn. However, it can be suggested that if litters are not adequately randomised, this may lead to biased conclusions depending on the extent of between litter variation.

Doolittle et al. (1976) suggested that maintenance of animals as litter-mate groups would keep the groups more uniform, even if the groups would vary in size. The results from the present study suggest that litters differ from each other in many parameters and if animals are not properly randomised, the probability for false positives will increase. Optimally litter-mates should be evenly allocated into all groups. In reality very few studies identify litter mates or allocate them into groups by litter. This cannot be applied when the number of groups exceeds the average number of one sex in a litter or when litter identification is lost.

Sex differences for several response variables are known among rats (*Festing, 1979*). In the present study, gender was expected to influence the physiological parameters; male rats grew faster and their adrenal glands were smaller than in female rats. Moreover, some of the serum parameters were affected by the gender. The influence of gender differences is quite well known and most commonly results in studies being restricted to a single sex of animals.

The concern of effects of enrichment procedures on study results is increasing (*Dahlborn et al., 1996; Howard, 1996; Eskola & Kaliste-Korhonen, 1998*). The optimum enrichment item is species appropriate (*Line, 1987*), engages animals in one or more beneficial ways (*Chmiel & Noonan, 1996*) and has no adverse effects on study results. Enrichment items in the present study did not have any real direct effects on the physiology of rats. Serum bilirubin levels seemed to be affected by the presence of items, but the post hoc-analyses did not reveal any statistical significancies. Based on these results, *i.e.* on absence of group mean changes, these enrichment items seem suitable for environmental enrichment in rats.

However, enrichment items seemed to have effects on variation, which does affect the number of animals needed. Refinement and reduction are two elements of the 3R-concept, but as seen from the results of the present study, sometimes there is conflict of interest. In fact there is a trade off - refining housing environment may necessitate

increased number of animals. To enrich or not should not be the question, but rather to know how to apply this piece of information in research.

In conclusion, the aspen enrichment items - especially the tube - seem to be species appropriate and suitable for rat enrichment since animals used these items and they did not have any adverse effects on rats. Furthermore, new materials were not introduced into the cage environment, which would further increase possible source of variation. However, even these enrichment items appear to have effects on the number of animals needed. This relationship should be emphasised in order to use both scientifically and ethically acceptable sample seizes.

Summary

Currently, environmental enrichment studies have focused on detection of differences in group means resulting from varied enrichments. The effects of enrichment procedures on variation and hence on study results have received far less attention. Within group variation in studies is directly related to adequate sample size and thus determines ethically correct number of animals to be used. The aim of this study was to evaluate the effects of enrichment items on clinical chemistry and physiological parameters of Wistar rats with the emphasis on the extent of variation. Forty-eight weaned female and male outbred Wistar rats from eight different litters were housed for five weeks in groups of four rats per cage with an aspen tube (20x12x12 cm), an aspen block (6x6x6 cm) or without enrichment items (control animals). At the end of the study, adrenals and brown adipose tissues were weighed, growth calculated from initial and final body weights and the following assays were performed from serum samples: AFOS, ALAT, ASAT, LDH, GGT, Pi, Ca, cholesterol, triglycerides, creatinine, total bilirubin, protein and corticosterone. Based on the variation detected, SOLO Power Analysis was used to calculate the smallest number of animals required (n) to detect an arbitrarily chosen 20 % difference in all means, when significance was set at $p=0.05$ and statistical power at 0.90. Based on the n-values, additional N-ratio ($n_{enrichment} / n_{control}$) was calculated to indicate, how many times more or less animals were needed in the enrichment group in comparison to the control group. The enrichment items did not have effects on group means. However, they seemed to influence the variation and hence on the number of animals needed, with enzyme parameters being the most sensitive in this respect. This relationship between enrichment, variation and number of animals needed should be kept in mind while designing experiments with ethically correct sample size.

References

Batchelor GR: The rest/activity rhythm of the laboratory rat housed under different systems. Anim. Technol. 1994, *45*, 181-187.

Beynen AC, MFW Festing, MAJ van Montfort: Design of animal experiments. In: van Zutphen, Baumans & Beynen (Eds.): Principles of Laboratory Animal Science. Elsevier, Amsterdam, 1993.

Blackshaw JK, DC Fenwick, AW Beattie, DJ Allan: The behaviour of chickens, mice and rats during euthanasia with chloroform, carbon dioxide and ether. Lab. Anim. 1988, *22*, 67-75.

Chamove AS: Cage design reduces emotionality in mice. Lab. Anim. 1989, *23*,215-219.

Chmiel DJ & M Noonan: Preference of laboratory rats for potentially enriching stimulus objects. Lab. Anim. 1996, *30*, 97-101.

Dahlborn K, BAA van Gils, IIA van de Weerd, JE van Dijk & V Baumans: Evaluation of long-term environmental enrichment in the mouse. Scand. J. Lab. Anim. Sci. 1996, *23*, 97-106.

Davies D: Preanalytical and analytical variation in animal clinical pathology: how to handle your samples. Abstracts. Scand-LAS symposium, Lillehammer, Norway 1998, 15-19 April.

Doolittle DP, SP Wilson & D Gieseking: Effect of caging on body weight and weight gain in mice. Lab. Anim. Sci. 1976, *26*, 556-561.

Erb HN: A statistical approach for calculating the minimum number of animals needed in research. ILAR News 1990, *32*, 11-16.

Eskola S & E Kaliste-Korhonen: Effects of cage type and gnawing blocks on weight gain, organ weights and open-field behaviour in

wistar rats. Scand. J. Lab. Anim. Sci. 1998, *25*, 180-193.

Eskola S, M Lauhikari, H-M Voipio & T Nevalainen: The use of aspen blocks and tubes to enrich the cage environment of laboratory rats. Scand. J. Lab. Anim. Sci. 1999, *26*,1-10.

Festing MFW: Inbred Strains in Biomedical Research. MacMillan Press, London, 1979.

Festing MFW: Are animal experiments well designed? Proceedings of the Fifth FELASA Symposium,1993, 8-11 June, Brighton UK.

Festing MFW: Reduction of animal use: experimental design and quality of experiments. Lab. Anim. 1994, *28*, 212-221.

Howard BR: Enrichment strategies. Proceedings of the sixth FELASA Symposium, Basel, Switzerland, 1996, 1-9.

Hubrecht R: Fundamentals of environmental enrichment. LASA Newsletter, Autumn issue: 1997, 8-9.

Kaliste-Korhonen E, S Eskola, T Rekilä & T Nevalainen: Effects of gnawing material, group size and cage level in rack on Wistar rats. Scand. J. Lab. Anim. Sci. 1995, *22*, 291-299.

Leard BL, RD Alsaker, WP Porter & LP Sobel: The effect of haemolysis on certain canine serum chemistry parameters. Lab. Anim. 1990, *24*, 32-35.

Line S: Environmental enrichment for laboratory primates. J. Am. Vet. Med. Ass. 1987, *190*, 854-859.

McGregor PK & SJ Ayling: Varied cages result in more aggression in male CFLP mice. Appl. Anim. Behav. Sci. 1990, *26*, 277-281.

Paton W: Man & Mouse. Animals in medical research. Oxford University Press, Oxford, 1984.

Russell WMS & RL Burch: The Principles of Humane Experimental Technique. Methuen & Co Ltd (Reprinted by UFAW, South Mimms, Potters Bar, 1992), London, 1959.

SOLO Statistical System Power Analysis. BMDP Statistical Software Inc., Los Angeles, CA, USA, 1991.

Sonntag O: Haemolysis as an interference factor in clinical chemistry. J. Clin. Chem. Clin. Biochem. 1986, *24*, 127-139.

Swanson HH & NE van de Poll: Effects of an isolated or enriched environment after handling on sexual maturation and behaviour in male and female rats. J. Reprod. Fert. 1983, *69*, 65-171.